AMENDMENTS TO THE CLAIMS

Upon entry of this amendment, the following listing of claims will replace all prior versions and listings of claims in the pending application.

IN THE CLAIMS

Please amend claims 1 and 8, and add claims 14-33 as follows:

1. (Currently Amended) A method for maximizing throughput while avoiding overload of one or more servers, comprising the steps of:

intercepting, via an interface unit, a client request for information from the server;

determining, by said interface unit, the current server performance, wherein said server performance is based on one or more of: the number of active connections opened to the server, the response time of the server, and the rate at which said response time is changing;

forwarding, by the interface unit, said client request to the server if said current server performance is below or within a range determined for optimal performance, whereby avoiding overload of the server; and

where server performance is outside such optimal range, buffering, by the interface unit, the client request until said current server performance is within the optimal range for server performance.

2. (Previously Presented) The method of claim 1, wherein said buffering utilizes a first-in-first-out method.

3. (Previously Presented) The method of claim 1, wherein said buffering comprises the steps of:

determining a preferred client value for said client request; and determining the position of said client request in a queue based on a preferred client value.

4. (Previously Presented) The method of claim 3, wherein said preferred client value may be partly determined by one or more of the network address (including either or both of the internet address and the port address) of said client request, by a header related to said client request, by

- 2 -

previous requests from the client of said client request, and by a cookie related to said client request.

5. (Previously Presented) The method of claim 1, further comprising the step of multiplexing connections to the server, whereby said multiplexed connections may be reused for different client requests.

6. (Previously Presented) The method of claim 1, further comprising the step of closing connections to the server as a way of reducing server load and improving server performance.

7. (Previously Presented) The method of claim 1, wherein the step of determining the current server performance may further be determined by the number of pending requests sent to the server and server error/overload messages from the server.

8. (Currently Amended) A system for maximizing throughput while avoiding overload of a server, comprising

    an interface unit for intercepting a client request for information from the server, wherein said interface unit determines the current server performance based on the number of connections opened to the server, the response time of the server, and the rate at which said response time is changing, wherein said interface unit forwards said client request to the server if said current server performance is below or within a range determined for optimal performance, whereby avoiding overload of the server, and where server performance is beyond such optimal range, ~~wherein~~ said interface unit buffers the client request until said current server performance is within the optimal range for server performance.

9. (Previously Presented) The system of claim 8, wherein said interface unit buffers said client request by utilizing a first-in-first-out method.

10. (Previously Presented) The system of claim 8, wherein said interface unit buffers said client request by determining a preferred client value for said client request and determining the position of said client request in a queue based on a preferred client value.

- 3 -

11. (Previously Presented) The system of claim 10, wherein said preferred client value may be partly determined by one or more of the network address (including either or both of the internet address and the port address) of said client request, by a header related to said client request, by previous requests from the client of said client request, and by a cookie related to said client request.

12. (Previously Presented) The system of claim 8, wherein said interface unit multiplexes connections to the server, whereby said multiplexed connections may be reused for client requests from different clients.

13. (Previously Presented) The system of claim 8, wherein said current server performance may further be determined by the number of pending requests sent to the server and server error/overload messages.

14. (New) A method for controlling a rate of establishing network connections to a server, the method comprising the steps of:

(a) intercepting, by an interface unit, a request from a client to establish a transport layer connection with a server, the interface unit providing a pool of at least a first transport layer connection;

(b) determining, by the interface unit, a number of transport layer connections established with the server; and

(c) opening, by the interface unit, a second transport layer connection to the server if the number of transport layer connections established with the server is below a maximum number of connections for the server.

15. (New) The method of claim 14, wherein step(c) further comprises opening the second transport layer connection if the first transport layer connection to the server is not free for use by the client.

- 4 -

16.   (New) The method of claim 14, comprising determining the maximum number of connections for the server based on one of a response time of the server or a rate at which the response time is changing.

17.   (New) The method of claim 14, comprising determining the maximum number of connections for the server based on one of a number of pending requests sent to the server or an error or overload messages from the server.

18.   (New) The method of claim 14, comprising if the first transport layer connection is available for use by the client, multiplexing, by the interface unit, via the first transport layer connection one or more requests from the client to access information on the server and one or more requests by a second client to access information on the server.

19.   (New) The method of claim 14, comprising queuing, by the interface unit, the request in a buffer until one of the first transport layer connection is available for use by the client or the second transport layer connection is opened.

20.   (New) The method of claim 19, wherein the interface unit utilizes a first-in-first-out method of queuing requests in the buffer.

21.   (New) The method of claim 19, wherein said queuing comprises the steps of:
    determining a preferred client value for the request of the client; and
    determining a position of the request in the buffer based on a preferred client value.

22.   (New) The method of claim 23, wherein the preferred client value is determined by one of a network address, a port address, a header or a cookie related to the request of the client.

23.   (New) The method of claim 14, comprising closing, by the interface unit, one of the transport layer connections to the server pooled by the interface unit.

- 5 -

24.   (New) A system for controlling a rate of establishing network connections to a server, the system comprising:

    means for intercepting, by an interface unit, a request from a client to establish a transport layer connection with a server, the interface unit providing a pool of at least a first transport layer connection to the server;

    means for determining, by the interface unit, a number of transport layer connections established with the server; and

    means for opening, by the interface unit, a second transport layer connection to the server if the number of transport layer connections established with the server is below a maximum number of transport layer connections for the server.

25.   (New) The system of claim 24, comprising means for opening the second transport layer connection if the first transport layer connection to the server is not free for use by the client.

26.   (New) The system of claim 24, comprising means for determining the maximum number of connections for the server based on one of a response time of the server or a rate at which the response time is changing.

27.   (New) The system of claim 24, comprising means for determining the maximum number of connections for the server based on one of a number of pending requests sent to the server or an error or overload messages from the server.

28.   (New) The system of claim 24, comprising means for multiplexing, by the interface unit, via the first transport layer connection, if the first transport layer connection is available for use by the client, one or more requests from the client to access information on the server and one or more requests by a second client to access information on the server.

29.   (New) The system of claim 24, comprising means for queuing, by the interface unit, the request in a buffer until one of the first transport layer connection is available for use by the client or the second transport layer connection is opened.

-6-

Application No. 09/912,401   Docket No. CTX-171

30. (New) The system of claim 29, wherein the interface unit utilizes a first-in-first-out method of queuing requests in the buffer.

31. (New) The system of claim 29, comprising means for determining a preferred client value for the request of the client; and determining a position of the request in the buffer based on a preferred client value.

32. (New) The system of claim 29, wherein the preferred client value is determined by one of a network address, a port address, a header or a cookie related to the request of the client.

33. (New) The system of claim 29, comprising means for closing, by the interface unit, one of the transport layer connections to the server pooled by the interface unit.

- 7 -